

08852684



05/09/97

[illegible]

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: PRESTIA, PAUL F.
- (B) REGISTRATION NUMBER: 23,031
- (C) REFERENCE/DOCKET NUMBER: GH-50008

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-407-0700
- (B) TELEFAX: 610-407-0701
- (C) TELEX: 846169

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3,881 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTTTGCGCCC	ACAAAATACA	CCGACGATGC	CCGATCTACT	TTAAGGGCTG	AAACCCACGG	60
GCCTGAGAGA	CTATAAGAGC	GTTCCCTACC	GCCATGGAAC	AACGGGGACA	GAACGCCCCG	120
GCCGCTTCGG	GGGCCCCGAA	AAGGCACGGC	CCAGGACCCA	GGGAGGCGCG	GGGAGCCAGG	180
CCTGGGCCCC	GGGTCCCCAA	GACCCCTGTG	CTCGTTGTCT	CCGCGGTCCT	GCTGTTGGTC	240
TCAGCTGAGT	CTGCTCTGAT	CACCCAACAA	GACCTAGCTC	CCCAGCAGAG	AGCGGCCCCA	300
CAACAAAAGA	GGTCCAGCCC	CTCAGAGGGA	TTGTGTCCAC	CTGGACACCA	TATCTCAGAA	360
GACGGTAGAG	ATTGCATCTC	CTGCAAATAT	GGACAGGACT	ATAGCACTCA	ATGGAATGAC	420
CTCCTTTTCT	GCTTTCGCTG	CACCAGGTGT	GATTTCAGGTG	AAGTGGAGCT	AAGTCCCTGC	480
ACCACGACCA	GAAACACAGT	GTGTCAAGTG	GAAGAAGGCA	CCTTCCGGGA	AGAAGATTCT	540
CCTGAGATGT	GCCGGAAGTG	CCGCACAGGG	TGTCCCAGAG	GGATGGTCAA	GGTCGGTGAT	600
TGTACACCCT	GGAGTGACAT	CGAATGTGTC	CACAAAGAAT	CAGGCATCAT	CATAGGAGTC	660
ACAGTTGCAG	CCGTAGTCTT	GATTGTGGCT	GTGTTTGTTC	GCAAGTCTTT	ACTGTGGAAG	720

AAAGTCCTTC	CTTACCTGAA	AGGCATCTGC	TCAGGTGGTG	GTGGGGACCC	TGAGCGTGTG	780
GACAGAAGCT	CACAACGACC	TGGGGCTGAG	GACAATGTCC	TCAATGAGAT	CGTGAGTATC	840
TTGCAGCCCA	CCCAGGTCCC	TGAGCAGGAA	ATGGAAGTCC	AGGAGCCAGC	AGAGCCAACA	900
GGTGTCAACA	TGTTGTCCCC	CGGGGAGTCA	GAGCATCTGC	TGGAACCGGC	AGAAGCTGAA	960
AGGTCTCAGA	GGAGGAGGCT	GCTGGTTCCA	GCAAATGAAG	GTGATCCCAC	TGAGACTCTG	1020
AGACAGTGCT	TCGATGACTT	TGCAGACTTG	GTGCCCTTTG	ACTCCTGGGA	GCCGCTCATG	1080
AGGAAGTTGG	GCCTCATGGA	CAATGAGATA	AAGGTGGCTA	AAGCTGAGGC	AGCGGGCCAC	1140
AGGGACACCT	TGTACACGAT	GCTGATAAAG	TGGGTCAACA	AAACCGGGCG	AGATGCCTCT	1200
GTCCACACCC	TGCTGGATGC	CTTGAGACG	CTGGGAGAGA	GACTTGCCAA	GCAGAAGATT	1260
GAGGACCACT	TGTTGAGCTC	TGGAAAGTTC	ATGTATCTAG	AAGGTAATGC	AGACTCTGCC	1320
ATGTCCTAAG	TGTGATTCTC	TTCAGGAAGT	CAGACCTTCC	CTGGTTTACC	TTTTTTCTGG	1380
AAAAAGCCCA	ACTGGACTCC	AGTCAGTAGG	AAAGTGCCAC	AATTGTCACA	TGACCGGTAC	1440
TGGAAGAAAC	TCTCCCATCC	AACATCACCC	AGTGGATGGA	ACATCCTGTA	ACTTTTCACT	1500
GCACTTGGCA	TTATTTTTAT	AAGCTGAATG	TGATAATAAG	GACACTATGG	AAATGTCTGG	1560
ATCAFTCCGT	TTGTGCGTAC	TTTGAGATTT	GGTTTGGGAT	GTCATTGTTT	TCACAGCACT	1620
TTTTTATCCT	AATGTAAATG	CTTTATTTAT	TTATTTGGGC	TACATTGTAA	GATCCATCTA	1680
CACAGTCGTT	GTCCGACTTC	ACTTGATACT	ATATGATATG	AACCTTTTTT	GGGTGGGGGG	1740
TGCGGGGCAG	TTCACTCTGT	CTCCCAGGCT	GGAGTGCAAT	GGTGCAATCT	TGGCTCACTA	1800
TAGCCTTGAC	CTCTCAGGCT	CAAGCGATTG	TCCCACCTCA	GCCATCCAAA	TAGCTGGGAC	1860
CACAGGTGTG	CACCACCACG	CCCGGCTAAT	TTTTTGTATT	TTGTCTAGAT	ATAGGGGCTC	1920
TCTATGTTGC	TCAGGGTGGT	CTCGAATTCC	TGGACTCAAG	CAGTCTGCCC	ACCTCAGACT	1980
CCCAAAGCGG	TGGAATTAGA	GGCGTGAGCC	CCCATGCTTG	GCCTTACCTT	TCTACTTTTA	2040
TAATTCGTGA	TGTTATTATT	TTATGAACAT	GAAGAACTT	TAGTAAATGT	ACTTGTTTAC	2100
ATAGTTATGT	GAATAGATTA	GATAAACATA	AAAGGAGGAG	ACATACAATG	GGGAAGAAG	2160
AAGAAFTCCC	CTGTAAGATG	TCACTGTCTG	GGTTCAGGCC	CTCCCTCAGA	TGTACTTTGG	2220
CTTCAATGAT	TGGCAACTTC	TACAGGGGCC	AGTCTTTTGA	ACTGGACAAC	CTTACAAGTA	2280
TATGAGTATT	ATTTATAGGT	AGTTGTTTAC	ATATGAGTCG	GGACCAAAGA	GAAGTGGATC	2340
CACGTGAAGT	CCTGTGTGTG	GCTGGTCCCT	ACCTGGGCAG	TCTCATTTGC	ACCCATAGCC	2400
CCCATCTATG	GACAGGCTGG	GACAGAGGCA	GATGGGTTAG	ATCACACATA	ACAATAGGGT	2460
CTATGTCATA	TCCCAAGTGA	ACTTGAGCCC	TGTTTGGGCT	CAGGAGATAG	AAGACAAAAT	2520
CTGTCTCCCC	ACGTCTGCCA	TGGCATCAAG	GGGAAGAGT	AGATGGTGCT	TGAGAATGGT	2580
GTGAAATGGT	TGCCATCTCA	GGAGTAGATG	GCCCCGGCTCA	CTTCTGGTTA	TCTGTCACCC	2640
TGAGCCCATG	AGCTGCCTTT	TAGGGTACAG	ATTGCCTACT	TGAGGACCTT	GGCCGCTCTG	2700
TAAGCATCTG	ACTCATCTCA	GAAATGTCAA	TTCTTAAACA	CTGTGGCAAC	AGGACCTAGA	2760
ATGGCTGACG	CATTAAGGTT	TTCTTCTTGT	GTCCTGTTCT	ATTATTGTTT	TAAGACCTCA	2820

GTAACCATTT CAGCCTCTTT CCAGCAAACC CTTCTCCATA GTATTTTCAGT CATGGAAGGA 2880  
 TCATTTATGC AGGTAGTCAT TCCAGGAGTT TTTGGTCTTT TCTGTCTCAA GGCATTGTGT 2940  
 GTTTTGTTCC GGGACTGGTT TGGGTGGGAC AAAGTTAGAA TTGCCTGAAG ATCACACATT 3000  
 CAGACTGTTG TGTCTGTGGA GTTTTAGGAG TGGGGGGTGA CCTTCTGGT CTTTGCACTT 3060  
 CCATCCTCTC CCACTTCCAT CTGGCATCCC CACGCGTTGT CCCCTGCACT TCTGGAAGGC 3120  
 ACAGGGTGCT GCTGCTTCCT GGTCTTTGCC TTTGCTGGGC CTTCTGTGCA GGACGCTCAG 3180  
 CCTCAGGGCT CAGAAGGTGC CAGTCCGGTC CCAGGTCCCT TGTCCTTCC ACAGAGGCCT 3240  
 TCCTAGAAGA TGCATCTAGA GTGTCAGCCT TATCAGTGTT TAAGATTTTT CTTTTATTTT 3300  
 TAATTTTTTT GAGACAGAAT CTCACTCTCT CGCCAGGCT GGAGTGCAAC GGTACGATCT 3360  
 TGGCTCAGTG CAACCTCCGC CTCCTGGGTT CAAGCGATTC TCGTGCCTCA GCCTCCGGAG 3420  
 TAGCTGGGAT TGCAGGCACC CGCCACCACG CCTGGCTAAT TTTTGTATTT TTAGTAGAGA 3480  
 CGGGGTTTCA CCATGTTGGT CAGGCTGGTC TCGAACTCCT GACCTCAGGT GATCCACNTT 3540  
 GGCCTCCGAA AGTGCTGGGA TATACAAGGC GTGAGCCACC AGCCAGGCCA AGATATTNTT 3600  
 NTAAAGNNAG CTTCCGGANG ACATGAAATA ANGGGGGGTT TTGTTGTTTA GTAACATTNG 3660  
 GCTTTGATAT ATCCCCAGGC CAAATNGCAN GNGACACAGG ACAGCCATAG TATAGTGTGT 3720  
 CACTCGTGGT TGGTGTCTT TCATGGTTCT GCCCTGTCAA AGGTCCCTAT TTGAAATGTG 3780  
 TTATAATACA AACAAGGAAG CACATTGTGT ACAAATACT TATGTATTTA TGAATCCATG 3840  
 ACCAAATTAA ATATGAAACC TTATATAAAA AAAAAAAAAA A 3881

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 411 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Gln Arg Gly Gln Asn Ala Pro Ala Ala Ser Gly Ala Arg Lys  
 1 5 10 15  
 Arg His Gly Pro Gly Pro Arg Glu Ala Arg Gly Ala Arg Pro Gly Pro  
 20 25 30

```

Arg Val Pro Lys Thr Leu Val Leu Val Val Ala Ala Val Leu Leu Leu
      35              40              45
Val Ser Ala Glu Ser Ala Leu Ile Thr Gln Gln Asp Leu Ala Pro Gln
      50              55              60
Gln Arg Ala Ala Pro Gln Gln Lys Arg Ser Ser Pro Ser Glu Gly Leu
      65              70              75              80
Cys Pro Pro Gly His His Ile Ser Glu Asp Gly Arg Asp Cys Ile Ser
              85              90              95
Cys Lys Tyr Gly Gln Asp Tyr Ser Thr Gln Trp Asn Asp Leu Leu Phe
              100              105              110
Cys Leu Arg Cys Thr Arg Cys Asp Ser Gly Glu Val Glu Leu Ser Pro
              115              120              125
Cys Thr Thr Thr Arg Asn Thr Val Cys Gln Cys Glu Glu Gly Thr Phe
              130              135              140
Arg Glu Glu Asp Ser Pro Glu Met Cys Arg Lys Cys Arg Thr Gly Cys
      145              150              155              160
Pro Arg Gly Met Val Lys Val Gly Asp Cys Thr Pro Trp Ser Asp Ile
              165              170              175
Glu Cys Val His Lys Glu Ser Gly Ile Ile Ile Gly Val Thr Val Ala
              180              185              190
Ala Val Val Leu Ile Val Ala Val Phe Val Cys Lys Ser Leu Leu Trp
              195              200              205
Lys Lys Val Leu Pro Tyr Leu Lys Gly Ile Cys Ser Gly Gly Gly Gly
      210              215              220
Asp Pro Glu Arg Val Asp Arg Ser Ser Gln Arg Pro Gly Ala Glu Asp
      225              230              235              240
Asn Val Leu Asn Glu Ile Val Ser Ile Leu Gln Pro Thr Gln Val Pro
              245              250              255
Glu Gln Glu Met Glu Val Gln Glu Pro Ala Glu Pro Thr Gly Val Asn
              260              265              270
Met Leu Ser Pro Gly Glu Ser Glu His Leu Leu Glu Pro Ala Glu Ala
      275              280              285
Glu Arg Ser Gln Arg Arg Arg Leu Leu Val Pro Ala Asn Glu Gly Asp
      290              295              300
Pro Thr Glu Thr Leu Arg Gln Cys Phe Asp Asp Phe Ala Asp Leu Val
      305              310              315              320
Pro Phe Asp Ser Trp Glu Pro Leu Met Arg Lys Leu Gly Leu Met Asp
              325              330              335
Asn Glu Ile Lys Val Ala Lys Ala Glu Ala Ala Gly His Arg Asp Thr
              340              345              350
Leu Tyr Thr Met Leu Ile Lys Trp Val Asn Lys Thr Gly Arg Asp Ala
      355              360              365
Ser Val His Thr Leu Leu Asp Ala Leu Glu Thr Leu Gly Glu Arg Leu
      370              375              380
Ala Lys Gln Lys Ile Glu Asp His Leu Leu Ser Ser Gly Lys Phe Met
      385              390              395              400
Tyr Leu Glu Gly Asn Ala Asp Ser Ala Met Ser End
              405              410 411

```

26050-135500

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1062 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGACCTCCT	TTTCTGCTTG	CGCTGCACCA	GGTGTGATTC	AGGTGAAGTG	GAGCTAAGTC	60
CCTGCACCAC	GACCAGAAAC	ACAGTGTGTC	AGTGCGAAGA	AGGCACCTTC	CGGGAAGAAG	120
ATTCTCCTGA	GATGTGCCGG	AAGTGCCGCA	CAGGGTGTCC	CAGAGGGGATG	GTCAAGGTCG	180
GTGATTGTAC	ACCCTGGAGT	GACATCGAAT	GTGTCCACAA	AGAATCAGGC	ATCATCATAG	240
GAGTCACAGT	TGCAGCCGTA	GTCTTGATTG	TGGCTGTGTT	TGTTTGCAAG	TCTTTACTGT	300
GGAAGAAAGT	CCTTCCTTAC	CTGAAAGGCA	TCTGCTCAGG	TGGTGGTGGG	GACCCCTGAGC	360
GTGTGGACAG	AAGCTCACAA	CGACCTGGGG	CTGAGGACAA	TGTCCTCAAT	GAGATCGTGA	420
GTATCTTGCA	GCCCACCCAG	GTCCCTGAGC	AGGAAATGGA	AGTCCAGGAG	CCAGCAGAGC	480
CAACAGGTGT	CAACATGTTG	TCCCCCGGGG	AGTCAGAGCA	TCTGCTGGAA	CCGGCAGAAG	540
CTGAAAGGTC	TCAGAGGAGG	AGGCTGCTGG	TTCCAGCAAA	TGAAGGTGAT	CCCACTGAGA	600
CTCTGAGACA	GTGCTTCGAT	GACTTTGCAG	ACTTGGTGCC	CTTTGACTCC	TGGGAGCCGC	660
TCATGAGGAA	GTTGGGCCTC	ATGGACAATG	AGATAAAGGT	GGCTAAAGCT	GAGGCAGCGG	720
GCCACAGGGA	CACCTTGTA	ACGATGCTGA	TAAAGTGGGT	CAACAAAACC	GGGCGAGATG	780
CCTCTGTCCA	CACCCTGCTG	GATGCCTTGG	AGACGCTGGG	AGAGAGACTT	GCCAAGCAGA	840
AGATTGAGGA	CCACTTGTTG	AGCTCTGGAA	AGTTCATGTA	TCTAGAAGGT	AATGCAGACT	900
CTGCCATGTC	CTAAGTGTGA	TTCTCTTCAG	GAAGTCAGAC	CTTCCCTGGT	TTACCTTTTT	960
TCTGGAAAAA	GCCCAACTGG	ACTCCAGTCA	GTAGGAAAGT	GCCACAATTG	TCACATGACC	1020
GGTACTGGAA	GAAACTCTCC	CATCCAACAT	CACCCAGTGG	AT		1062

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 303 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asp	Leu	Leu	Phe	Cys	Leu	Arg	Cys	Thr	Arg	Cys	Asp	Ser	Gly	Glu	Val
1				5				10					15		
Glu	Leu	Ser	Pro	Cys	Thr	Thr	Thr	Arg	Asn	Thr	Val	Cys	Gln	Cys	Glu
			20					25				30			
Glu	Gly	Thr	Phe	Arg	Glu	Glu	Asp	Ser	Pro	Glu	Met	Cys	Arg	Lys	Cys

7